
[W] [O] [R] [E] [S] [E] [I] [N] [G] (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPorch_PP protein - protein database search, using Smith-Waterman algorithm
Run on: Sat May 13 10:28:20 2000; Maspar time 4.63 Seconds
337.619 Million cell updates/sec
Tabular output not generated.

Title: >US-09-331-631-5
Description: (145-210) from US09331631.pep (3 of 4)
Perfect Score: 514
Sequence: 1 KRDPQOREYEDCRHCEQOE.....PQSGSGRYEEGEEKSDNP 66

Scoring table: PAM 150
Gap 11

Sequences: 188963 seqs, 23686106 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq35
1:geneseqp

Statistics: Mean 25.287; Variance 98.331; scale 0.257

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description	Pred. No.
1	514	100.0	625	1	Macadamia integrifolia	1.26e-43
2	481	93.6	666	1	Macadamia integrifolia	1.25e-40
3	452	87.9	666	1	Macadamia integrifolia	5.25e-37
4	162	31.5	525	1	Theobroma cacao antiml	4.96e-07
5	162	31.5	566	1	Sequence encoded by 67	4.96e-07
6	141	27.4	590	1	Gossypium hirsutum anti	5.22e-05
7	120	23.3	637	1	Hordeum vulgare antiml	4.91e-03
8	88	17.1	919	1	Human androgen recepto	3.59e+00
9	84	16.3	388	1	Myobacterium species	7.88e+00
10	84	16.3	593	1	Zea mays antimicrobial	7.88e+00
11	82	16.0	918	1	Human androgen recepto	1.16e+01
12	82	16.0	919	1	Androgen receptor	1.16e+01
13	82	16.0	919	1	Human androgen recepto	1.16e+01
14	79	15.4	112	1	Myobacterium species	2.07e+01
15	79	15.4	126	1	Myobacterium species	2.07e+01
16	78	15.2	271	1	Parapox virus F9L prot	2.51e+01
17	78	15.2	303	1	Dirofilaria immitis pa	2.51e+01
18	78	15.2	376	1	Sequence of protein H.	2.51e+01
19	78	15.2	376	1	Sequence of protein H.	2.51e+01
20	77	15.0	86	1	GST-HD fusion protein	3.04e+01
21	77	15.0	86	1	GST-HD fusion protein	3.04e+01
22	77	15.0	395	1	Mouse SRY-related prot	3.04e+01
23	77	15.0	2289	1	Protein derived from r	3.04e+01

ID	Score	Query Match	Length DB	ID	Description	Pred. No.
24	76	14.8	662	1	R83304	Canine distemper virus
25	75	14.6	400	1	W17831	Human mevalonate pyrop
26	75	14.6	425	1	R13792	E75B exon B1 polypepti
27	75	14.6	537	1	R75188	Osteoinductive retrovi
28	75	14.6	669	1	W37483	Mouse liver cancer-ori
29	75	14.6	816	1	R71111	Spinochordella ataxia
30	75	14.6	2482	1	W23996	Human mitotin amino ac
31	75	14.6	2482	1	R72826	Kinetochore protein CE
32	75	14.6	3248	1	R97925	Epitope tagged TBP pro
33	74	14.4	371	1	W73369	Antimicrobial maie pe
34	73	14.2	35	1	R21079	G. max truncated SBP1
35	73	14.2	444	1	W90340	G. max SBP1 protein
36	73	14.2	524	1	W90339	Stenocarpus sinuatus a
37	72	14.0	28	1	W62841	Zea mays antimicrobial
38	72	14.0	33	1	W62836	Moloney murine leukemi
39	72	14.0	263	1	R80533	G. max truncated SBP2
40	72	14.0	409	1	W90342	G. max SBP2 protein
41	72	14.0	489	1	W90341	Candida albicans CAC1A
42	72	14.0	971	1	W48896	Actin-filament binding
43	72	14.0	1829	1	Y07242	Human secreted protein
44	71	13.8	205	1	W78191	Rat vascular smooth mu
45	71	13.8	807	1	Y04264	

ALIGNMENTS

RESULT 1
ID W62830 standard; Protein; 625 AA.
AC W62830: (first entry)
DE 27-OCT-1998
DT Macadamia integrifolia antimicrobial protein.
KW antimicrobial protein; infestation; control.
OS Macadamia integrifolia.
FH Key Location/Qualifiers
FT Peptide 1..28 /note="signal peptide"
FT Protein 29..666 /note="mature protein"
FT W09827805-A1.
PD 02-JUL-1998.
PN 22-DEC-1997; AU0874.
PR 20-DEC-1996; AU-004275.
PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
DR WPI: 98-377279/32.
DR N-PSDB: VA2316.
PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT useful for controlling microbial infestations of plants or mammals
PS Claim 1; Page 43-45; 96pp; English.
CC The sequence is that of an antimicrobial protein which can
CC be used to control microbial infestations in plants and mammalian
CC animals.
CC Sequence 625 AA:
SQ
Query Match 100.0%; Score 514; DB 1; Length 625;
Best Local Similarity 100.0%; Pred. No. 1.26e-43;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 145 KRDPQOREYEDCRHCEQOEPRLYQOCRCOQORORHGRGDLNPNQSGSGRYEEGEE 204
QY 145 KRDPQOREYEDCRHCEQOEPRLYQOCRCOQORORHGRGDLNPNQSGSGRYEEGEE 204
Db 205 KOSDNP 210
QY 205 KOSDNP 210
RESULT 2
ID W62829 standard; Protein; 666 AA.
AC W62829:
DE 27-OCT-1998 (first entry)
DT Macadamia integrifolia antimicrobial protein.
KW antimicrobial protein; infestation; control.

QY	150	QREEDCSYHC-EQGE-PRLOYSQKQCEQQRQNRHNGSGDLMNPRQSGSGNYEEGEEKQS	207
Db	141	NNP	143
QY	208	DNP	210

RESULT 6
ID W62832 standard; Protein; 590 AA.
AC W62832:
DT 27-OCT-1998 (first entry)
DE Gossypium hirsutum antimicrobial protein.
KW antimicrobial protein; infestation; control.
OS Gossypium hirsutum.
PN MO9827805-A1.
PD 02-JUL-1998.
PE 22-DEC-1997: AU08874.
PR 20-DEC-1996: AU-004275.
PA (REFR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PI Bower NJ, Goulter KC, Green JL, Manners JH, Marcus JP;
DR WI: 98-377279/32.
PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
UT useful for controlling microbial infestations of plants or mammals
PS Claim 1, Page 49-51: 96pp. English.
CC The sequence is that of an antimicrobial protein which can
CC be used to control microbial infestations in plants and mammalian
CC animals.
SQ Sequence 590 AA;

RESULT 7
ID W62837 standard; Protein; 637 AA.
AC W62837;
DT 27-OCT-1998 (first entry)
DE Hordeum vulgare antimicrobial protein.
KW antimicrobial protein; infestation; control.
OS Hordeum vulgare.
PN W09827805-A1.
PD 02-JUL-1998.
PE 22-DEC-1997: AU0874.
PR 20-DEC-1996: AU-004275.
PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOG.
PI Bower NI, Goullier KC, Green JL, Manners JM, Marcus JP;
DR WP: 98-377279/32.
PT Novel anti-microbial protein from e.g. *Macadamia integrifolia* -
PT useful for controlling microbial infestations of plants or mammals
PS Claim 1: Page 60-62; 96pp: English.
CC The sequence is that of an antimicrobial protein which can
CC be used to control microbial infestations in plants and mammalian
CC animals.
CC Sequence 637 AA;
IQ

RESULT	8
ID	P90996 standard; protein; 919 AA
AC	P90996;
DT	28-FEB-1990 (first entry)

DE	Human androgen receptor DNA clone.
KM	Androgen receptor; TR2 polypeptide.
OS	Homo sapiens.
FH	Key
FT	location/qualifiers
FT	region
FT	1..919
FT	/tag= a
FT	/product=98 kd polypeptide
FT	185..919
FT	/tag= b
FT	/product=79 kd polypeptide
PN	WC8909223-A.
PD	05-OCT-1989.
PE	24-MAR-1989; U01238.
PR	30-MAR-1988; US-176107.
PA	(ARCH-) Arch Development Corp.
PI	Liao S, Chang C;
DRI	WPI; 89-309501/42.
DR	N-PSDB; N91577.
PT	New DNA encoding new androgen receptor and TR2 polypeptide(s) - able to bind DNA, and derived antibodies, useful for receptor assay and purification.
PT	Purification.
PS	Claim 8; fig. 3; 60pp; English.
CC	The protein is used to raise antibodies for receptor assays and for affinity purification.
CC	The 98 kd product starts at the first Met codon; the 79 kd product starts from the second.
Q	Sequence 919 AA.

RESULT 9 standard; Protein; 388 AA.
ID Y04998 standard; Protein; 388 AA.
AC Y04998:
DT 06-JUL-1999 (first entry)
DE Mycobacterium species protein sequence 50B.
KW Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
KM hybridisation; detection; vaccine; immunisation; infection.
OS Mycobacterium sp.
PN W09090186-A2.
PD 25-FEB-1999.
PE 14-AUG-1998; F01813.
PR 11-SEP-1997; FR-011325.
RA 14-AUG-1997; FR-010404.
PA (INSP) INST PASTEUR.
PI Gicquel B, Lim EM, Pelicic V, Portnoi D, Gouquet de la Salmoniere Y,
PI Guigueno A;
DR WPI; 99-181045/15.
DR N-PSDB; X34249.
PT Mycobacterial DNA vectors containing reporter constructs - for
PT identifying coding or promoter sequences involved in
PT infection-associated protein expression
PS Claim 32; Fig 50B; 309pp; French.
CC Sequences Y04742-Y05000 and Y07201-Y07204 represent secreted proteins
CC from various Mycobacterium species microorganisms. The encoding
CC nucleotide sequences can be used as primers and probes for methods
CC for detecting and identifying mycobacteria, especially belonging to
CC the M. tuberculosis complex. The encoded proteins can be used in
CC vaccines for immunisation against a bacterial or viral infection.
QO Sequence 388 AA; #

145 KRDPQREYEDCRRCRCEQDEPRLQYQCRRCQEQQR 180

RESULT 10
ID W62835 standard; Protein: 593 AA.
AC W62835:
DT 27-OCT-1998 (first entry)
DE Zea mays antimicrobial protein.
KW antimicrobial protein; infestation; control.
OS zea mays.
PN W09827805-A1.
PD 02-JUL-1998.
PF 22-DEC-1997.
PR 20-DEC-1996; AU-004275.
PI (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PT Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
DR WPI: 98-377279/32.
PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
PI useful for controlling microbial infestations of plants or mammals
PS Claim 1: Page 58-60; 96pp; English.
CC The sequence is that of an antimicrobial protein which can
CC be used to control microbial infestations in plants and mammalian
CC animals.
SQ Sequence 593 AA;

Query Match 16.3%; Score 84; DB 1; Length 593;
Best Local Similarity 27.6%; Pred. No. 7,88e+00;
Matches 16; Conservative 20; Mismatches 17; Indels 5; Gaps 4;

39 OCVRCEDR-PWHRPCLEOCREERERKORSHRHEAD-DRSGSGSEDEREQEK 94

155 DCRRCCEQDEPRLQY-QCRRCQEQQR--HGKGDLMNPQSGSGRYEEGEEKOSDN 209

RESULT 11
ID R12223 standard; Protein: 918 AA.
AC R12223:
DT 20-AUG-1991 (first entry)
DE Human androgen receptor.
KW hAR; DNA-binding protein; steroid hormone.
OS Homo sapiens.
FH Key Location/Qualifiers
FT domain 556..626
FT /label=DNA-binding domain
FT /note="cysteine-rich"

MO9107423-A.
PD 30-MAY-1991.
PF 19-OCT-1990; U06015.
PR 17-NOV-1989; US-438775.
PA (ARCH-) ARCH DEV CORP.
PI Liao S, Chang C;
DR WPI: 91-178048/24.
DR N-PSDB: 012001.
PT Androgen receptor and TR2 DNA binding proteins - DNA sequences
PT and antibodies for detection and quantification methods
PS Claim 25: Fig 3; 79pp; English.
CC This sequence was deduced from a cDNA clone isolated by screening
CC commercially available human testis and prostate lambda gt11 cDNA
CC libraries. The sequence is very similar to that of rat AR and in
CC the DNA-binding domain it is identical to that of rAR DNA-binding
CC domain. Homology comparisons with other known steroid receptors
CC indicate that hAR is more closely related to glucocorticoid,
CC mineralo-corticoid and progesterone receptors than to v-erb-A or to
CC receptors for oestrogen, vitamin D and thyroid hormones.
SQ Sequence 918 AA;

Query Match 16.0%; Score 82; DB 1; Length 918;
Best Local Similarity 32.0%; Pred. No. 1.16e+01;
Matches 16; Conservative 11; Mismatches 23; Indels 0; Gaps 0;

64 QQQQQQQQQQETSPRQQQQQGGEDGSPQAHRRGPTGYLVDEEQPSOP 113
161 EQQEPRLQYQCRRCQEQQRHGKGDLMNPQSGSGRYEEGEEKOSDN 210

RESULT 12
ID W14783 standard; Protein: 919 AA.
AC W14783:
DT 22-JUN-1997 (first entry)
DE Androgen receptor.
KW Androgen receptor; acidic fibroblast growth factor; aFGF;
KW antisense; benign prostatic hyperplasia; prostate cancer; therapy.
OS Homo sapiens.
PN M09711170-A1.
PD 27-MAR-1997.
PF 20-SEP-1996; U15081.
PR 20-SEP-1995; US-004018.
PI (MORC-) WORCESTER FOUND BIOMEDICAL RES.
PT Zamecnik PA;
DR WPI: 97-202879/18.
DR N-PSDB: T63407.

PT Oligonucleotide(s) antisense to human androgen receptor and acidic
PT FGF genes - used to inhibit gene expression, for the treatment of
PT benign prostatic hyperplasia
PS Disclosure; Page 22-28; 51pp; English.
CC Human androgen receptor (W14783) binds testosterone and, acting
CC at the transcriptional level, regulates the growth of normal
CC prostatic cells. Antisense oligonucleotides (see also T63700,
CC T63404-05) based on an androgen receptor cDNA clone (see also
CC T63407) can be used to prevent androgen receptor gene expression,
CC thereby inhibiting the growth or survival of prostatic cells for
CC the treatment of benign prostatic hyperplasia and prostate cancer.
SQ Sequence 919 AA;

Query Match 16.0%; Score 82; DB 1; Length 919;
Best Local Similarity 32.0%; Pred. No. 1.16e+01;
Matches 16; Conservative 11; Mismatches 23; Indels 0; Gaps 0;

68 QQQQQQQQQQETSPRQQQQQGGEDGSPQAHRRGPTGYLVDEEQPSOP 117
161 EQQEPRLQYQCRRCQEQQRHGKGDLMNPQSGSGRYEEGEEKOSDN 210

RESULT 13
ID P93109 standard; Protein: 919 AA.
AC P93109:
DT 19-MAR-1990 (first entry)
DE Human androgen receptor.
KW Human androgen receptor; ployclonal antibody; cancer.
OS Homo sapiens.
PN M08909791-A.
PD 19-OCT-1989.
PF 13-APR-1989; U01548.
PR 14-APR-1988; US-182646.
PA (UNC-) University of North Carolina.
PI French FS, Wilson EM, Joseph DR, Lubahn DB;
DR WPI: 89-324206/44.
DR N-PSDB: N91772.
PT DNA encoding androgen receptor protein - useful for transforming
PT eukaryotic hosts for protein expression and subsequent antibody prodn.
PS Disclosure; Fig. 4; 44pp; English.
CC Androgen receptor protein (AR) is used to produce mono- or poly-clonal
CC antibodies. These are used for the detection and quantification of AR in
CC the presence of endogenous androgen, as androgen will not interfere with
CC binding. They may be used in assays to determine and quantify cellular
CC distribution of AR in tumour tissue, and are esp. useful for evaluating
CC prostate cancers to determine responsiveness to androgen withdrawal
CC therapy.
SQ Sequence 919 AA;

Query Match 16.0%; Score 82; DB 1; Length 919;
Best Local Similarity 32.0%; Pred. No. 1.16e+01;
Matches 16; Conservative 11; Mismatches 23; Indels 0; Gaps 0;

68 QQQQQQQQQQETSPRQQQQQGGEDGSPQAHRRGPTGYLVDEEQPSOP 117

OY 161 EQQEPRLQYOCORRCOEQRQHGRCGDLMPNQRGSGRIECCXKOSDP 210

RESULT 14
ID Y04866 standard; Protein; 112 AA.

AC Y04866:
DT 06-JUL-1999 (first entry)
DE Mycobacterium species protein sequence 19C'.
KW Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
OS Mycobacterium sp.
PN W09909186-A2.
PD 25-FEB-1999.
PF 14-AUG-1998; F01813.
PR 11-SEP-1997; FR-011325.
PR 14-AUG-1997; FR-010404.
PA (INSP) INST PASTEUR.
PI Gicquel B, Lim EM, Pelicic V, Portnoi D, Goguet de la Salmoniere Y,
PI Guigueno A;
DR WPI: 99-181045/15.
DR N-PSDB: X34118.
PT Mycobacterial DNA vectors containing reporter constructs - for
PT Identifying coding or promoter sequences involved in
PT Infection-associated protein expression
PS Claim 32; Fig 19C; 309pp; French.
CC Sequences Y04742-Y05000 and Y07201-Y07204 represent secreted proteins
CC from various Mycobacterium species microorganisms. The encoding
CC nucleotide sequences can be used as primers and probes for methods
CC for detecting and identifying mycobacteria, especially belonging to
CC the M. tuberculosis complex. The encoded proteins can be used in
CC vaccines for immunisation against a bacterial or viral infection.
SQ Sequence 112 AA;

Query Match 15.4%; Score 79; DB 1; Length 112;
Best Local Similarity 32.6%; Pred. No. 2.07e+01;

Matches 14; Conservative 8; Mismatches 19; Indels 2; Gaps 1;

Db 25 RRLCGKHTAQORFXCANPGLRSRVQGRKRGDPRRQHRGEG 67
OY 157 RRHCEQEPRLQYOCORRCOEQRQHGRCGDLMPNQRGSG 197

RESULT 15
ID Y04861 standard; Protein; 126 AA.

AC Y04861:
DT 06-JUL-1999 (first entry)
DE Mycobacterium species protein sequence 19A.
KW Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
OS Mycobacterium sp.
PN W09909186-A2.
PD 25-FEB-1999.
PF 14-AUG-1998; F01813.
PR 11-SEP-1997; FR-011325.
PR 14-AUG-1997; FR-010404.
PA (INSP) INST PASTEUR.
PI Gicquel B, Lim EM, Pelicic V, Portnoi D, Goguet de la Salmoniere Y,
PI Guigueno A;
DR WPI: 99-181045/15.
DR N-PSDB: X34113.
PT Mycobacterial DNA vectors containing reporter constructs - for
PT Identifying coding or promoter sequences involved in
PT Infection-associated protein expression
PS Claim 32; Fig 19A; 309pp; French.
CC Sequences Y04742-Y05000 and Y07201-Y07204 represent secreted proteins
CC from various Mycobacterium species microorganisms. The encoding
CC nucleotide sequences can be used as primers and probes for methods
CC for detecting and identifying mycobacteria, especially belonging to
CC the M. tuberculosis complex. The encoded proteins can be used in
CC vaccines for immunisation against a bacterial or viral infection.
SQ Sequence 126 AA;

Query Match 15.4%; Score 79; DB 1; Length 126;

Best Local Similarity 32.6%; Pred. No. 2.07e+01;
Matches 14; Conservative 8; Mismatches 19; Indels 2; Gaps 1;

Db 39 RRLCGKHTAQORFXCANPGLRSRVQGRKRGDPRRQHRGEG 81
OY 157 RRHCEQEPRLQYOCORRCOEQRQHGRCGDLMPNQRGSG 197

Search completed: Sat May 13 10:28:29 2000
Job time : 9 secs.

